Comprehension and Situation Awareness

Francis T. Durso
Texas Tech University, USA
Katherine A. Rawson
Kent State University, USA
and
Sara Girotto
Texas Tech University, USA

The central role of comprehension in human cognition has been recognized by both basic and applied researchers. Basic research in comprehension, conducted under the rubrics of text comprehension, language processing, and reading research, is a large part of the field of psycholinguistics in particular and cognitive science in general. Applied research in comprehension, conducted under the rubric of situation awareness (SA), is a large part of the field of cognitive ergonomics in particular and human factors in general.

Although SA has (many) very specific definitions in the literature (e.g., Endsley 1990; see Rousseau *et al.* 2004), in this chapter we think of SA as comprehension, or understanding, of a dynamic environment. In fact, the origin of the term SA in aviation highlights comprehension: the component of tactical flight operations which involves the pilot's understanding. There are some advantages to this way of thinking. The term *comprehension* carries less baggage than does the term *awareness*. For example, comprehension allows for implicit (e.g., Croft *et al.* 2004) as well as explicit (Gugerty 1997) information. It may also help applied researchers sidestep semantic entanglements, like SA as product and SA as process. Finally, the term *comprehension* acknowledges the connections to the large basic research database on reading comprehension from which SA work has benefited, and invites continued comparisons between comprehension of dynamic situations and comprehension of text.

Lack of understanding when performing complex cognitive tasks can have dramatic consequences (see Casey 1993; Chiles 2002). The incident at Three Mile Island was a result of operators draining coolant because they misunderstood the coolant level to be too high. As another example, nearly 5000 people died between 1978 and 1992 in flights

Handbook of Applied Cognition: Second Edition. Edited by Francis T. Durso.

that were under control but flown into terrain (CFIT: controlled flight into terrain). According to Woodhouse and Woodhouse (1995), 74 per cent of these CFIT accidents were due to a lack of awareness on the part of the flight crew, as opposed to non-adherence or proficiency/skill failure. Similarly, Durso *et al.* (1998b) reported that 62 per cent of the operational errors while controlling flights between airports (en route) were made by air traffic controllers unaware that a loss of separation was developing. In many domains, superior performance is linked to superior SA, and not to other, less cognitive skills. For example, Horswill and McKenna (2004) argue that of all the components of driving the *only* one that correlates with safety is SA, and not for example vehicle control skills.

Currently, there is no clear consensus on how operators understand the dynamic environment in which they work. There are efforts to understand SA from a variety of perspectives. Although it is possible to look at SA from perspectives other than cognitive information-processing (cf. Adams *et al.* 1995), there is certainly reason to treat SA as a cognitive construct (Endsley & Bolstad 1994; Endsley 1995; Durso & Gronlund 1999). As examples, Carretta *et al.* (1996) asked which abilities and personality traits predicted SA in F-15 pilots. They used cognitive measures of working memory, spatial ability, time estimation, and perceptual speed. Psychomotor skills and the Big Five personality traits were also assessed. SA was determined by judgments of peers and supervisors using a Likert scale ranging from 1 (acceptable) to 6 (outstanding). When flying experience was controlled, general cognitive ability was found to predict SA, but not psychomotor skills or personality traits. Similarly, O'Hare (1997), using an SA measure requiring pilots to "scan multiple information sources, evaluate alternatives, establish priorities, and select and work on the task that has the highest priority at the moment" (WOMBAT; Roscoe 1997), was able to discriminate elite pilots from ordinary pilots. Chaparro *et al.* (1999) showed that drivers' ability to recognize hazards was dependent on their divided and selective attention ability.

Endsley and colleagues have championed a general cognitive approach in the literature. In 1995, Endsley sketched a framework of the cognitive processes likely to underlie SA. In this chapter, we add to that work by borrowing from the reading comprehension literature to specify more precisely at least one possible sequence of processes that gives rise to understanding of dynamic situations. As in the reading comprehension literature, in this chapter *comprehension* is not viewed as one stage in a serial sequence of stages. To us comprehension is the phenomenon that emerges from an orchestra of cognitive processes. Perception is involved in all comprehension, but, perhaps less obviously, so are top-down, predictive processes, and bottom-up event integrating processes. Thus, although we break down the comprehension process into constituents for the sake of exposition, all component processes operate in a highly integrated fashion.

Thinking of SA as something like reading comprehension, but in a dynamic environment, is the analogy that drives this chapter. Thus, unlike the previous coverage of SA in the first edition of this book (Durso & Gronlund 1999), which let the extant literature direct the shape of the review, this chapter takes a somewhat more top-down approach. Our intent here is to review the literature on SA, but to do so in the context of a model of situation comprehension. By so doing, we explore the appropriateness of an analogy with reading comprehension for furthering understanding of SA, and review the literature that has accumulated since the Durso and Gronlund review of SA. Like all analogies, there are limits, but comparing something poorly understood (e.g., SA) to something well

understood (e.g., reading comprehension), has been useful in every scientific domain (e.g., Oppenheimer 1956).

MEASURING COMPREHENSION

We begin our analogy with methodology. Because basic and applied investigators have differed in their epistemic goals, their methods have also differed. Basic cognitive research explains comprehension by characterizing the underlying cognitive mechanisms involved. In order to do this, researchers have developed particular methodologies that are used in relatively simple, controlled experiments. Because those experiments are about language, the domain is almost universally static text. The empirical facts that have emerged from this work have allowed researchers to reach a surprising level of consensus on the theoretical underpinnings of comprehension. Of course, particular models differ in their specifics, but researchers interested in comprehension in reading agree on much.

Measuring Text Comprehension

A wide variety of measures has been developed to measure text comprehension. The kinds of measures used in text comprehension research may be roughly grouped into three categories, as shown in Table 7.1.

Self-report Measures

Self-report measures include those in which individuals are asked to report their subjective beliefs about their comprehension (metacomprehension judgments) or report on their own thinking during comprehension (verbal protocols). Understanding individuals' beliefs about and conscious experiences of their own comprehension may be important for understanding their subsequent judgments and behaviors. However, metacomprehension judgments and verbal protocols are limited in the extent to which they can be used as measures of comprehension and in the extent to which they reveal the nature of underlying cognitive processes.

Accuracy Measures

The second category includes objective measures of comprehension in which accuracy is the primary dependent variable of interest. These objective measures can be further divided into those that primarily measure memory for text content and those that primarily measure deeper comprehension (Kintsch 1994), although no measure provides a "pure" assessment of either memory or comprehension. Note this important distinction between memory and comprehension – an individual could memorize this paragraph well enough to recite it without really understanding the concepts or ideas in it. Conversely, most of our readers will understand this paragraph quite well while reading, although they may not be able to remember many of the specific details afterwards.

Table 7.1 Kinds of measures used in text comprehension research

Task	Descriptions and examples	Sample references
Self-report measures		
Metacomprehension judgments	Judging how well a text has been understood Predicting how well one will do on a test Estimating how well one has done on a test	Rawson *et al.* (2002) Thiede and Anderson (2003) Maki (1998) Rawson and Dunlosky (2002) Maki *et al.* (1990) Maki *et al.* (1994)
Verbal protocols	Unconstrained "thinking out loud" while reading Answering open-ended questions while reading (e.g., after each sentence of a narrative, explaining why the event described therein happened)	Kendeou and van den Broek (2005) Suh and Trabasso (1993) Magliano *et al.* (1999)
Performance accuracy measures		
Memory measures		
Recall	Free recall (e.g., "Write down everything you can remember from the text you just read") Cued recall (e.g., "What is the definition of ________ ?")	Kintsch (1998) Rawson and Kintsch (2004) Myers *et al.* (1987)
Recognition	e.g., "Which of the following sentences appeared in the text you just read?" e.g., "Which item below is the definition of ________ ?"	Zwaan (1994) Kintsch *et al.* (1990)
Comprehension measures		
Inference questions	Forming a connection between two ideas that was not explicitly stated in the text, forming a connection between an idea and relevant prior knowledge, or drawing valid conclusions from ideas stated in the text	Mayer *et al.* (1996) Mayer and Jackson (2005)
Transfer problems	Applying principles from one domain to new problems in another domain	Rawson and Kintsch (2005)
Concept organization	Drawing concept maps or diagrams, concept sorting, similarity judgments	McNamara and Kintsch (1996)

Table 7.1 *Continued*

Task	Descriptions and examples	Sample references
Performance time measures		
Self-paced reading times	Computer presents units of material (words, phrases, sentences, or paragraphs) one at a time and individual advances through units at own pace; amount of time spent reading each unit is recorded. Reading times in different experimental conditions re often compared. Correlations between reading times and variables of theoretical interest are often computed	Graesser *et al.* (1980) Millis *et al.* (1998)
Eye movements	Intact text material is presented and eye tracking equipment records (a) amount of time spent looking at each region and (b) the pattern of eye movement between regions	Just and Carpenter (1980) Rayner (1998) Wiley and Rayner (2000)
Implicit query Lexical decision Naming Recognition	Reading is interrupted at target locations by a probe presented for speeded response. These measures are typically used to estimate the activation level of a target concept, based on the extent to which response times to the target are faster than to a control. Decide whether a string of letters forms a word Pronounce a word as fast as possible	Long *et al.* (1992) Klin *et al.* (1999) Wiley *et al.* (2001) McKoon *et al.* (1996) Singer *et al.* (1992) O'Brien and Albrecht (1991)
Explicit query	Indicate whether a word appeared in the text just read Reading is interrupted at target locations by a question requiring a speeded response (e.g., the sentence "Bob believed Bill because he was gullible" is followed by the query, "Who was gullible?")	

In most studies using accuracy measures, individuals read an entire text and later answer questions based on the text material. These measures are useful for understanding the product readers acquire from a text and the extent to which this information can be retained and used subsequently. However, these measures are limited for investigating the nature of the processes involved during reading. For example, if an individual cannot answer an inference question after reading a text, it does not mean that inferential processing was not taking place during reading (the inference may have been computed but then forgotten). Likewise, the ability to answer the question does not necessarily mean that the inference was made during reading (the inference may have been made at time of test).

Latency Measures

The third category includes measures in which latency is the primary dependent variable of interest. In contrast to the first two kinds of measure, these tasks are accepted as more useful for studying the nature of the underlying cognitive processes and the mental representations involved during reading because the measures are taken at the time of processing (vs. measuring the product after process has been completed, i.e., the representation that is still available after the reading task). For example, to explore whether individuals make predictive inferences while reading, Klin *et al.* (1999) presented readers with short stories in one of two versions, one that was consistent with a target predictive inference and one that was not. To illustrate, one story described the protagonist either as having lost his job or as having been given a healthy raise. The story then goes on to say that the protagonist really wanted to give his wife something special for her birthday, and he noticed a beautiful ruby ring sitting unattended on a department store counter. The story ended with "He quietly made his way closer to the counter." Immediately after the last sentence, the word STEAL appeared on the screen and participants simply had to say the word aloud as quickly as possible. People were faster to say the word when the man had been described as losing his job than when he had been described as getting a raise. Presumably, readers had already activated the concept steal in the former case because they predicted that he was going to steal the ring from the counter (vs. buy it in the latter case).

One measure of time that has been particularly informative in reading research is fixation duration or dwell time. Key assumptions have allowed eye movement research in reading to have powerful implications for underlying cognition (Just & Carpenter 1980). According to the *immediacy assumption*, interpretation of a stimulus (e.g., a word) begins as soon as the stimulus is fixated. According to the *eye–mind assumption*, "the eye remains fixated on a word as long as the word is being processed . . . there is no appreciable lag between what is being fixated and what is being processed" (Just & Carpenter 1980, pp. 330–331).

In summary, performance time measures such as these can support fine-grained analysis of the mental representations and processes involved during comprehension. In the text-comprehension literature, the adoption of one performance time measure over another often depends upon weighing tradeoffs between task intrusiveness that may disrupt or alter normal comprehension processing versus the ease of interpreting the data that are acquired. When possible, the use of converging methods is routinely recommended (e.g., Klin *et al.* 1999).

Measuring Comprehension of Dynamic Environments

A variety of different measures has been developed to measure SA. One scheme (Durso & Gronlund 1999) classifies measures into three general types: subjective measures, query methods, and implicit performance measures. Such a classification can be productively compared with a classification that comes from our analogy to reading comprehension: self-reports, accuracy, and time (see Table 7.2).

Subjective Measures

Subjective measures, as in reading research, typically require the operator to make self-judgments about understanding. One of the most well-known subjective measures is the Situation Awareness Rating Technique (SART) developed by Taylor (1990), which requires operators to make judgments along a number of dimensions, some of which capture impressions of workload, whereas others capture more cognitive dimensions.

As with reading comprehension, subjective measures of SA are metacomprehension, not comprehension, measures. However, this does not mean that subjective measures are not useful. In most industrial situations, it is very important that objective SA and subjective judgments of SA coincide. In a future where the operator must decide whether to turn on an intelligent aid, or in the present where the operator must admit he needs help, research on *meta-SA* is needed. Nevertheless, these measures reveal nothing about the underlying processes of comprehension.

Unlike in reading research, we also find a few subjective measures based on *observer* reports. Unlike reading, in dynamic environments controlled by operators, it is at least possible that behaviors signal the level of situation comprehension. Efforts to formalize such experiences have asked subject matter experts (SMEs) to observe the operator's performance and then to rate the participant's level of SA; the Carretta *et al.* (1996) study with F-15 pilots discussed earlier is one example. See also SA/BARS (Neal *et al.* 1998).

Accuracy Measures

By analogy to the reading comprehension literature, SA measures using accuracy as a dependent variable tend to measure the product of situation awareness, that is, the final product of the situation comprehension processes; for example, location of own ship, awareness of the mode of the aircraft, and navigational awareness are SA products in aviation. SA accuracy measures include query methods and implicit performance measures. Query methods explicitly ask the operator to report a piece of task-relevant information (see Jeannot *et al.* 2003 for a review). Implicit performance measures examine how an operator responds to an SA-revealing event embedded, either naturally or by clever experimenters, into the scenario. It is of fundamental importance that the tasks used as implicit performance measures of SA can be performed successfully by an operator with good SA but unsuccessfully by one with poor SA. Similarly, event detection (Gugerty & Falzetta 2005) requires the operator to detect particular embedded events, like swerves or decelerations.

Table 7.2 Measures of SA comparing a classification scheme from the literature with a classification drawn from the analogy to reading comprehension. One classification scheme reported in the SA literature

		Subjective	Query	Implicit performance
Classification scheme of text comprehension methodologies	Self-reports	SART (Taylor 1990); SAPS (Deighton 1997; Jensen 1999); C-SAS (Dennehy 1997); SASHA_Q (Jeannot et al. 2003); SA-SWORD (Vidulich & Hughes 1991); SARS (Waag & Houck 1994); Verbal protocols (Metalis 1993)		
	Accuracy		SAGAT (Endsley 1990); SALSA (Hauβ *et al.* 2000, 2001); SAPS (Deighton 1997; Jensen 1999); Explicit probes (Vidulich *et al.* 1994)	Andre *et al.* (1991) Measures of Effectiveness (Vidulich *et al.* 1994); Implicit probes (Vidulich et al. 1994) Gugerty and Falzetta (2005)
	Time		SPAM (Durso *et al.* 1995; 1998); SASHA_L (Jeannot *et al.* 2003)	Busquets *et al.* (1994)

As Table 7.2 suggests, most objective measures of SA, whether query methods or implicit performance methods, rely on accuracy. Examining accuracy is valuable for understanding which types of information about a certain situation the operator retains. The most widely used query method, the Situation Awareness Global Assessment Technique (SAGAT), developed by Endsley (1990), relies on accuracy. In order to administer this measure (see Jones & Kaber 2004), the experimenter prepares a series of questions relevant to the task the operator will have to perform. The simulation is stopped at points, and all the information relevant to the task is physically removed from the operator who, at that point, is asked to answer the questions previously prepared. In addition to the possible effects that the intrusiveness of the method might cause, SAGAT has been criticized for relying too heavily on conscious memory (e.g., Sarter & Woods 1991; Durso *et al.* 1998a). The criticisms raised by reading comprehension researchers apply here as well: If the operator does not have a good picture of the situation when queried, that does not mean that she did not have the picture while performing the task. A common example is highway amnesia. Although the driver may wonder if she stopped at the traffic light for which she has no memory, failure to answer correctly does not mean she had an SA failure at the traffic light. Underwood *et al.* (2002) showed that experienced drivers can be inaccurate in judgments about what they looked at just moments ago. The contrary can hold as well: An operator may form a mental image at the point of query that may differ from the one actually present during task performance.

Latency Measures

Methods that use time as a dependent variable are a step toward investigating the cognitive processes that underlie situation comprehension. Accuracy tells us about SA only when it fails; response time has the potential to help us in investigating what happens when SA succeeds. This logic led to the use of response time to understand human memory and opened research into semantic memory and knowledge structures (Lachman *et al.* 1979).

Thus far, only a few measures have been developed to measure SA using response time, although many can certainly be adapted. Some are implicit performance measures. For example, Busquets *et al.* (1994) had participants land on one runway while another aircraft was to land on another runway. Occasionally, the second aircraft would deviate and try to land on the first runway. The time to take action to avoid the second aircraft was the implicit performance measure of SA.

Query methods collecting response time have also been developed. In the Situation-Present Assessment Method (SPAM; Durso *et al.* 1998a; Durso & Dattel 2004), the operator is given unsolicited requests for information (hence the unusual acronym) while he or she is performing the task. For example, the operator can be asked which one of two airplanes has the lower altitude or, given the current speed, which one will reach a waypoint first. The logic of SPAM is that if the information is immediately available to the operator, response time to the query should be short. If the information is not available, but the operator knows where to find the information, then response time will be longer, but not as long as the case in which the operator does not know where to find the information. Thus, SPAM leaves the operator in context and assumes that knowing where to find a piece of information could be indicative of good SA, even if the information was not available in memory. In fact, if a piece of information was immediately available in the environment, it might be a poor idea to use limited resources to remember it. Other details including how to eliminate workload effects and how to construct the queries can be found in Durso and Dattel (2004). SPAM has recently spawned other measures (Jeannot *et al.* 2003).

SPAM has been shown to improve prediction over a large battery of psychological tests (Durso *et al.* in press). Performance on an ATC simulator was predictable from various cognitive and occasionally noncognitive variables. Of importance here was the finding that SPAM improved predictability of handoff delay times and air traffic errors above and beyond the battery of standard tests, but off-line queries using accuracy did not. Thus, an SA measure like SPAM was able to capture additional variance in performance.

Finally, for fixation duration or dwell time, it is not clear how, or even if, the eye-movement assumptions from text comprehension, hold in a dynamic environment. Rather than process a stimulus and then move on to the next without revisiting the stimulus as does a skilled reader, the skilled industrial operator makes many small fixations (about 2.5/s), revisiting displays for very brief periods (Moray 1990). In the literature there are cases in which experts have longer fixations (Williams & Davids 1998), cases in which their fixations are shorter (Crundall & Underwood 1998), and cases in which no differences in fixation durations are found (Helsen & Pauwels 1993; Williams *et al.* 2002) between experts and novices. Perhaps the best way to understand the difficulty in equating real-world dwell times and cognitive processing is to imagine where you look when driving in the country vs. in the city. Chapman and Underwood (1998) showed that drivers fixate longer on a rural road than an urban one. It seems unlikely that more cognitive processing is occurring on the rural road, and thus researchers must be careful in interpreting eye fixations in real-world dynamic environments, where the operator has choices and where the task is uncontrolled.

Nevertheless, from more controlled, yet dynamic, situations there comes hope that fixation duration may reveal insights into SA. For example, when encountering a dangerous situation, both experts and novice drivers increase fixation durations (Chapman & Underwood 1998). There have even been findings in accord with some of the more subtle discoveries in reading research. For example, when soccer players did not anticipate correctly, they tended to fixate on the player with the ball longer than during correct trials (Helsen & Starkes 1999). This result mirrors nicely those found in reading research when expectations are violated and the offending information receives a longer fixation.

In summary, meta-SA (i.e., self-report) measures are valuable in revealing metacognitive assumptions the operator is making when controlling a dynamic environment, but they tell us little about either the product of comprehension or the processes. SA accuracy measures can tell us about the product of comprehension. Adding latency measures allow insights into the processes as well. Although SA and reading comprehension measures need not agree, they did show a number of similarities. Researchers in both fields must keep several important dimensions in mind when selecting measures, including temporal proximity, availability of external information, invasiveness, and congruence between the SA measure and the performance measure. For example, measures will be more informative about the process of comprehension to the extent that they are temporally proximal to task performance, whereas measures that are taken after task completion are more likely to reflect the products of comprehension. Finally, and perhaps most exciting, additional SA measures can be developed by adapting other reading comprehension measures (Table 7.1) to dynamic environments.

TOWARD A MODEL OF COMPREHENSION OF DYNAMIC SITUATIONS

In this chapter, we use key assumptions from theories of text comprehension to motivate the development of a model of situation comprehension. To foreshadow, an important starting assumption is that text comprehension is not one process. Rather, comprehension is best thought of as a system of cognitive processes. These processes operate together in a coordinated fashion to encode and integrate various kinds of information. In fact, the most widely accepted theoretical claim in text comprehension research is that comprehension processes operate at three basic levels, with different kinds of information encoded at each level (Kintsch 1998). The *surface level* representation includes the exact words and grammatical structures used to form the sentences. In contrast to the linguistic information encoded at the surface level, the *textbase* contains semantic information. That is, the textbase is the representation of the meaning that is extracted from the linguistic input. Finally, like the textbase, the *situation model* also contains meaningful information. However, whereas the textbase primarily includes information that is explicitly stated in the text, the situation model integrates the textbase with prior world knowledge to form a fuller representation of the situation being described in the text. A great deal of research has been dedicated to exploring the processes involved in encoding and integration at each of these levels.

Using these theoretical assumptions of text comprehension models as a foundation, we propose a model of the comprehension of dynamic environments. A schematic of the model appears in Figure 7.1. Each component of the model is described below. To overview, according to our model, situation comprehension involves several different cognitive processes that encode and integrate various kinds of information. By analogy to text comprehension, situation comprehension involves a *surface level* representation that includes the objects in the environment and the structural relationships between them (i.e., scenes). The *eventbase* includes the semantic information that is extracted from the perceptual input. Finally, the *situation model* integrates the semantic information that can be derived from the external input with prior knowledge to form a fuller representation of the situation.

An Illustration

For illustrative purposes, consider an individual who is talking on a cell phone while driving. If she is particularly engaged in the phone conversation, she may fail to check her rearview mirror, an attentional failure at the surface level that results in inadequate sampling of information from the scene behind her. She may encode coarse-grained information about the presence of several cars in the lanes ahead, she may fail to encode the cascade of brake lights on the cars in front of her as they near an intersection.

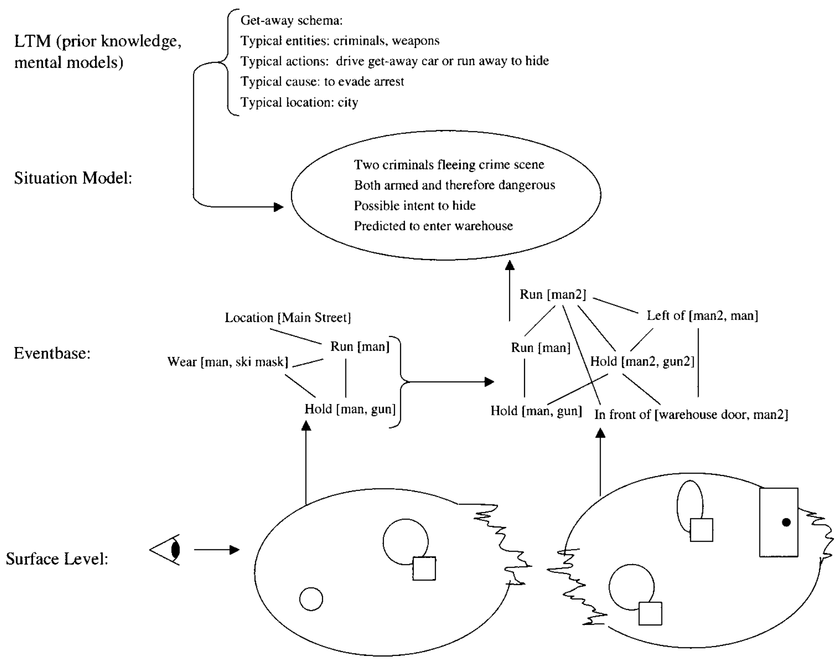

Figure 7.1 A schematic of the model. At the surface level, the operator processes perceptual information: objects and scenes. Information from this level is added to the eventbase, which operates on the basis of promiscuous activation and pruning. A skilled operator will have a mental model containing relevant information, like causation, relevant to the prototypical situation. The mental model, together with the actual situation, instantiates a situation model. Details of the model are contained in the text

Even if adequate surface-level information is processed, the extraction of semantic information may be inadequate, and thus our distracted driver may also have deficient eventbase processing. While sitting at the intersection, she may encode the perceptual features of the red light going off and a green arrow coming on, but fail to process the meaning of the green light (until the annoyed driver behind her sounds his horn). Similarly, as she approaches the four-way stop at the next intersection, she may encode the scene to her left in which a red sedan has just come to a stop. She may also encode the scene to her right in which a blue minivan rolls to a stop about the same time she does. However, what she may fail to do is to integrate the spatial and temporal information from the two scenes and to extract the relevant semantic information – namely, the information that determines who has right of way and should enter the intersection first.

Finally, our driver may have adequate surface and eventbase processing but may still have inadequate situation comprehension if the situation model fails. For example, our driver may process that the driver of an oncoming SUV appears to be oriented toward a child in the backseat. She may also integrate the spatial and temporal information from this scene with information from the other scenes she has sampled, such as that the other three vehicles (hers included) are reaching the intersection before the SUV. She may even successfully extract semantic information, such as encoding from the driver's facial features that he is yelling angrily at the child in the backseat. But if her situation model processing is incomplete, she may fail to grasp the unfolding situation – i.e., although the man in the SUV does not have right of way, he is not attending to the intersection and is not slowing down, and thus our driver should infer the likely outcome that the SUV will run the stop sign. If she fails fully to represent the situation in this way, she may act based only on her eventbase representation (i.e., she has the right of way) thus initiating what will turn out to be a disastrous left turn.

The Surface Level: Encoding Objects and Scenes

The surface level of a text representation includes the particular words and grammatical structures included in each sentence. At first glance, the analogy between text comprehension and situation comprehension may not seem useful, because dynamic environments obviously have different kinds of information than static linguistic information. However, at a broader level, the analogy invites us to ask what the words and sentences of a dynamic environment are. Dynamic environments contain objects and structural relationships between them that must be encoded (much like words and the grammatical relations between them). Thus, as a sentence comprises words and their syntactic relationships, we can think of a scene as a group of characteristic objects (De Graef *et al.* 1990; Bar & Ullman 1996).

According to Endsley's (1995) definition of SA, perceptual processing of individual units is the first step in situation assessment. Many studies have demonstrated that loss of SA is often caused by faulty perception. Jones and Endsley (1996) analyzed 262 pilot SA errors derived from 143 accidents. They found that 72 per cent of SA errors were due to perceptual and attentional processing errors. Durso *et al.* (1998b) reported that 50 per cent of the operational errors caused by air traffic controllers who were unaware that the error was occurring had perceptual or attentional errors underlying them. About 44 per cent of knives carried in baggage are not detected when the airport baggage screener fails to fixate directly on the knife's image, an attentional failure. Even when fixated, 15 per cent of the knives went undetected (McCarley *et al.* 2004). Thus, both attention to and perception of units in the environment are critical to SA.

Of course, the surface level involves not just the encoding of individual objects, but also the encoding of scenes. Scene identification is very quick, sometimes as fast as the identification of a single constituent object (Potter 1976; Friedman 1979; Biederman *et al.* 1982). In many cases, a single fixation can be sufficient to get the "gist" of a scene (Renninger & Malik 2004). For example, Schyns and Oliva (1994; Oliva & Schyns 1997) found that scenes could be identified when pictures were presented for only 45–135 ms.

Operators extract various cues from the environment in order to perform their task, whether it is predicting where the tennis serve will land or whether the approaching aircraft is hostile. These cues can be scenes, objects, or parts of objects. For example, Schyns and Oliva (1994) found that scenes could be identified from holistic cues like low-spatial frequency that preserves the spatial relations between large-scale structures existing in the scene without presenting the visual details needed to identify the individual objects.

Renninger and Malik (2004) found that humans confuse scenes with similar textures and that texture alone was able to account for correct categorization on eight of the ten scene categories presented. Thus, scene level information, independent from information relative to individual objects, can support scene identification (Biederman 1981, 1988; Schyns & Oliva 1994; Oliva & Schyns 2000).

Surface level processing has been shown to be of fundamental importance in hazard perception, or "situation awareness for dangerous situations in the traffic environment" (Horswill & McKenna 2004, p. 155). In typical research, participants watch a monitor and press a button when the situation presents a danger. Reaction time and accuracy both correlate with on-road evaluations of driving instructors (Mills *et al.* 1998). Experts detect hazards faster than novices (McKenna & Crick 1991; Avila & Moreno 2003), in part because experienced drivers know potential hazard locations (see Underwood *et al.*, Chapter 15, this volume). There is some evidence that conceptual categorization of cues (i.e., that is a category member) can be as rapid as perception of the simple presence of the cue (Secrist & Hartman 1993).

Superior SA can result not only if an operator detects a cue more quickly, but also if the cue is more diagnostic (Salas *et al.* 2001). Expert squash players seem to rely at least partially on perceptual extraction of better cues. For example, looking at response latency, Howarth *et al.* (1984) found that experts used information extracted prior to ball contact, whereas less skilled players relied on early ball-flight information. Some of this prior-to-contact information seems to be opponent movement (Abernethy *et al.* 2001) and some seem to be proximal cues (Abernethy & Russell 1987). These results have been replicated in a variety of sports (see Abernethy *et al.*, Chapter 13, this volume).

Clearly, failures to perceive the cues that indicate a danger can travel down the cognitive stream and lead to poor decisions and poor performance. Consider the fact that pilots sometimes decide to fly into storms, often with fatal consequences. Ineffective decision-making related to weather has led researchers to try to identify the psychological reasons why pilots decide to continue a flight when weather conditions are deteriorating (Wiggins & O'Hare 2003). Faulty perceptual classification seems at least partially to blame.

Wiggins and O'Hare (2003) developed a training program for novice pilots to facilitate identification of cues helpful in recognizing dangerous, weather-related situations. One strategy that is promising in teaching operators how to recognize relevant cues in the environment is cognitive apprenticeship (Druckman & Bjork 1991): Trainees work closely with experts on a series of activities that take place in the real environment.

Studies investigating the effectiveness of sport-specific perceptual training found that players improve their performance after they learn to use visual cues (James & Patrick 2004). Experiments (Williams *et al.* 2002; Farrow & Abernethy 2003) in which tennis players were trained with either implicit or explicit cue recognition revealed better performance of the two groups with respect to a control and a placebo group.

In summary, the literature has clear support that perception is an important component of situation comprehension. Perception in a dynamic environment can proceed by sampling and identifying objects and scenes, with the latter often occurring as quickly as the former. Activation of a scene can proceed from a characteristic constituent object or from holistic cues. Once a scene is identified, top-down influences on identification of constituent objects become possible. The perceptual components of SA, such as speed of scene identification and the relationship between scenes and objects, have important consequences for application. Training programs that focus on surface level processing have had success.

The Eventbase: Integrating Sequences of Meaningful Events

The Textbase in Text Comprehension

Whereas the surface level representation in text comprehension includes linguistic information (i.e., the exact words and grammatical structures in a sentence), the textbase represents the semantic information that is derived from the linguistic input, including the concepts denoted by words and the ideas denoted by the structural relationships between them. For example, "The man was bitten by the dog" and "The dog bit the man" have different grammatical structures but express the same idea or *proposition*.

Importantly, in addition to the concepts and propositions themselves, the textbase represents the semantic relationships between them. Such connections can be based on several different dimensions, including reference, causality, time, and space. For example, two propositions could be connected if they express events that occur in the same timeframe or spatial location. Likewise, a connection between two propositions may be represented when they refer to the same entity. The important point is that the representation of concepts and propositions alone is not enough. These elements must also be connected to one another, or integrated, to form a coherent representation. Furthermore, not only must connections between elements within a sentence be represented, but also connections between the elements in different sentences must be represented for a coherent representation of the text. Quite simply, not to represent relations between elements that are explicitly stated or strongly implied is to incompletely understand a text, and thus incoherent representations may lead to comprehension failures. Accordingly, an important issue concerns how connections between elements are formed. According to the construction-integration (CI) theory of comprehension (Kintsch 1988, 1998), only a limited amount of text material can be processed at one time due to limited cognitive capacity. Text comprehension thus proceeds in cycles, with the input in a given cycle roughly equivalent to a sentence. Each cycle involves two phases of processing. In the construction phase, representational "nodes" are created that correspond to the concepts and propositions extracted from the linguistic input. Each node can then activate associated concepts, propositions, and higher-order knowledge structures (e.g., schemata) from long-term memory, which are also included as nodes in the developing network. The other key process involved during the construction phase involves the formation of connections between nodes, based on various factors (e.g., time, space, reference, and causality).

The *integration* phase involves the spreading of activation throughout the network. Highly interconnected nodes will accumulate activation, whereas less well-connected nodes will lose activation and may drop from the network altogether. As a result of the spreading activation process, network nodes will vary in their activation level at the end of integration. Given the limited capacity of the processing system, the entire network cannot be carried over to the next processing cycle because some capacity must be available for the processing of the next input. According to the CI model, the network that remains after integration is stored in long-term memory but only the subset of nodes with the highest ending activation remain in working memory to participate in the next processing cycle. It is this "carrying over" of nodes from one cycle to the next that allows the integration of information across segments of text. Thus, an important part of processing is determining what to carry over from one processing cycle to the next – the success of integrating information across cycles will depend upon which nodes are carried over. The highly activated nodes in one cycle will usually, but not always, represent the information that is most related to the next input. If not, the representation will be incoherent (i.e., one form of comprehension failure) in the absence of additional processing.

Empirical Evidence from Text Comprehension Research

Previous research has provided support for each of these theoretical claims. Several studies have reported evidence for the representation of concepts and propositions (e.g., Kintsch *et al.* 1975; Murphy 1984; O'Brien *et al.* 1997). For example, Murphy (1984) showed that processing a word that introduces a new concept is more time-consuming than processing the same word when it refers to an existing concept. Studies involving multiple regression analyses have shown monotonic increases in sentence reading times, with each additional word introducing a new concept and with each additional proposition in a sentence, after controlling for other related variables (e.g., Graesser *et al.* 1980; Haberlandt *et al.* 1980; Graesser & Bertus 1998; Millis *et al.* 1998).

Research has also provided support for assumptions about the nature of the processes involved in constructing the textbase. For example, concerning the claim that input is processed in cycles, studies using reading time and eye movement measures have reported robust *wrap-up effects*. The wrap-up effect refers to the finding that reading times are substantially longer for the final word of a major clause or sentence than for non-boundary words (e.g., Just & Carpenter 1980; Haberlandt *et al.* 1986; Rayner *et al.* 2000; but see Magliano *et al.* 1993). These effects are attributed to the integration process, which presumably takes place at these boundaries. Some research has focused on investigating how the elements to be carried over from one processing cycle to the next are selected (e.g., Fletcher 1981; Malt 1985; Glanzer & Nolan 1986; Fletcher *et al.* 1990; McKoon *et al.* 1993). Computational models that simulate text processing based on the principles of the CI theory have also been successful at predicting the probability with which humans recall particular propositions from a text (e.g., Goldman & Varma 1995; Kintsch 1998; Rawson & Kintsch 2004), which provides converging evidence for the plausibility of the hypothesized processes.

The Eventbase in Situation Comprehension

By analogy to this work on textbase representations and processes in text comprehension, we posit that a complete representation of a dynamic environment involves the construction of an *eventbase*. Whereas the surface level representation in situation comprehension includes perceptual information (e.g., the objects and the spatial relationships between them in a scene), the eventbase represents the semantic information that is derived from the perceptual input.

Sometimes this information is quite distinct from the perceptual input. For example, an air traffic controller in the tower cab may observe the physical American 767 ascending from Runway 31 to 3 000 feet underneath a nearby Delta 757 that is descending to 3 000 feet, whereas an air traffic controller in the approach control (TRACON) may receive the corresponding information from a radar screen. Although the perceptual input is quite different in the two cases, the two air traffic controllers will likely develop a similar eventbase representation (e.g., each including a "concept" node for Plane A and a "concept" node for Plane B, a "proposition" expressing that Plane A is ascending, a "proposition" expressing that Plane B is descending, and the connections between these elements based on space, time, and reference). Previous applied research has acknowledged that a visual form can differ from a representational form (e.g., Woods 1991). On other occasions, the information in the eventbase is more dependent on the perceptual input. For example, Garsoffky *et al.* (2002) provide data that suggest that viewpoint is retained in the eventbase. Participants watched short (less than 30-second) clips of soccer goals and after each clip made yes/no recognition judgments on video stills. Regardless of level of expertise or from where in the clip the still was taken, recognition hits showed viewpoint dependency. Thus, when witnessing an event, the relationship between perception and semantics is less arbitrary than what would be expected from a strict analogy to reading comprehension. However, even when witnessing an event, the eventbase cannot be equivalent to the raw surface level type of information, nor can it be composed entirely of perceptual information.

According to the analogy, only a limited amount of the situation can be processed at one time due to limited cognitive capacity. Moray (1990) discusses how limited the processing is and how features of the environment allow experienced operators to deal with these limits. Industrial operators sample their environment about 2.5 times a second. If the environment is well structured, then these brief samples can take advantage of redundancies in the world. If it is unstructured, or the operator does not have the experience to take advantage of the structure, then situation awareness will be impaired. A favorite example is that when free-flying moths detect the ultrasonic pulse of a predatory bat, they fly a random flight path (Roeder 1962); such evasive actions make the predictability of the system low, thus limiting the bat's SA. Further, if the bandwidth of the environment is too high, that is if the environment is changing too rapidly, then processing will be beyond the operators' brief samples and again SA will suffer. In most tasks the bandwidth is acceptable (e.g., transportation) or divided up into teams (e.g., Unoccupied Aerial Vehicles) or otherwise reduced, although there are cases like low-altitude combat flying (Haber & Haber 2003) in which the operator is given explicit training on managing the bandwidth by explicitly learning times required to perform tasks as function of altitude.

Importantly, the eventbase also integrates semantic information across scenes. Consider a fighter pilot who encodes perceptual information from the instrument panel including a radar screen indicating two aircraft to the right, one at 45° and one at 120°. The pilot then looks through the right window and physically observes one aircraft flying slightly ahead. Each scene is encoded in one cycle of processing, much as each cycle of processing during text comprehension involves roughly one sentence. However, the information encoded from the first scene (the instrument panel) has yet to be integrated with the information encoded from the second scene (the view through the window). Perceptual information per se will not allow the pilot to make the referential connection between the physical object on the radar screen and the physical object observed through the window, because the two are quite dissimilar perceptually. However, the connection can be established at the semantic level. From the first scene, the pilot may represent a "concept" for Plane A, a "concept" for Plane B, a "proposition" expressing that Plane A is ahead of own ship, and a "proposition" expressing that Plane B is behind own ship. If these elements are carried over to the next cycle of processing, a referential link may be established with the "proposition" from the second scene expressing that a plane is ahead of our pilot's plane.

The analogy leads to the hypothesis that the processes involved in construction of an eventbase are similar to those involved in the construction of a textbase. Processing proceeds in cycles, with one scene processed in each cycle (e.g., a driver looking ahead to a stoplight, checking the rearview mirror, and then looking out the side window). Nodes will be created that correspond to the meaning of the objects in the scene as well as nodes expressing the semantic relationships between them. Each node may activate associated information from long-term memory. Connections will be formed between nodes based on various dimensions (e.g., time, space, reference), and then activation will be spread through the network resulting in the pruning of some nodes. Only those elements that are most highly active at the end of integration will be carried over to participate in the processing of the next scene.

To discriminate cleanly between the eventbase, which is not influenced by domain specific knowledge, and the situation model, which is, it is important to consider the ability to track changes outside of an area of expertise. Exactly this kind of work was begun in the 1960s by Yntema and colleagues (Yntema & Mueser 1960, 1962; Yntema 1963). For example, in the study by Yntema and Mueser (1960), participants saw an 8×8 grid of "doors," with a row representing an object (e.g., object "K") and a column representing an attribute (e.g., shape) that could take on four states (e.g., circle, square, triangle, heart). Messages read to the participant indicated the value that an attribute had and would continue to have until further notice. Occasionally, the procedure was interrupted to interrogate the participant about the present state of one of the variables (e.g., "What is the current shape of object K?"). Yntema and Mueser varied the number of objects with the same single attribute to be monitored and the number of attributes for the same, single object. Accuracy decreased as the number of attributes whose states were to be remembered increased. Monitoring one attribute across different objects was more difficult than monitoring multiple attributes for one object. For example, keeping track of the shape of eight objects was more difficult than keeping track of eight attributes of one object. Propositionally, this might be represented as SHAPE (K, circle), SHAPE (D, square), SHAPE (N, triangle) and so on, versus SHAPE (K, circle), FOOD (K, toast), WEATHER (K, stormy), and so on. These two eventbases could be represented as in Figure 7.2.

The eventbase representations in Figure 7.2 suggest why the Yntema results obtain. Consider how the processing cycles for these two sets of descriptions might proceed. For the left-hand set, the first processing cycle would involve construction of a concept node for "K," a proposition denoting that K was a circle, and the integration of these two nodes. These two nodes would be carried over for inclusion in the next cycle. However, because no connection can be established between these nodes and the subsequent input, they will likely be dropped during the spreading activation process of the next cycle. At a minimum, the more recent input will be more strongly activated and will thus be selected for carryover to the next processing cycle. Either way, the K nodes will not participate in any additional processing cycles. By comparison, consider the processing cycles for the right-hand set. The initial processing cycle would be the same, with construction of the two K nodes

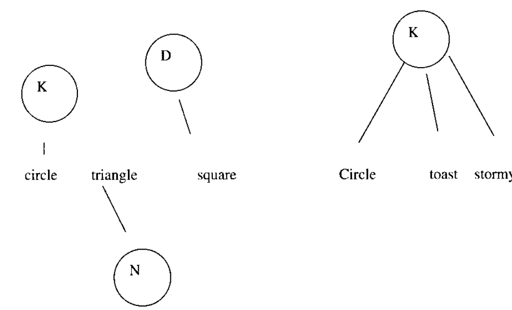

Figure 7.2 Eventbase representation of the monitoring task described in Yntema & Mueser (1960)

and then carryover of those nodes to the next cycle. However, in this case, both nodes can be connected to the nodes constructed for the subsequent input based on argument overlap (i.e., all propositions contain the argument "K"). Thus, the two original nodes are less likely to be dropped from the network during spreading activation. Additionally, even though not all four nodes will be carried over to the next cycle, those nodes that are carried over will probabilistically reactivate the other K nodes with which they are now connected during the construction phase, providing those K nodes with yet another opportunity to become linked to other K nodes and to accrue more activation. Importantly, the CI theory states that the retrievability of any given node is partly a function of the activation it accrues during processing (which increases with the number of processing cycles in which it participates) and the connection strength between that node and the retrieval cue. Both of these factors favor the latter set of descriptions, which provides an explanation for why recall is greater in this condition.

Integrating across scenes and eye fixations to create an eventbase is, in some ways, easier to appreciate in dynamic environments than in reading static text. Scan paths have been studied for years. It is well known that during the early phases of learning to scan, the scan paths are likely to be more variable (Wikman *et al.* 1998) but less flexible (Crundall & Underwood 1998). Recent analyses in some domains are becoming quite intricate. Consider driving. Underwood *et al.* (2003) identified scanpaths of differing number of fixations. Given that experienced drivers have better SA than less experienced ones, it is unlikely that the advantage is due to simple fixations since Underwood *et al.*'s drivers were comparable in that regard. Fixations were heavily dependent on the immediate past for experienced drivers, but regardless of where the novice driver was looking at fixation N, he or she looked at the road far ahead on fixation N + 1. Even though the two groups experienced the same input (comparable single fixations) the information carried from one processing cycle to another (different scan paths) created different eventbases, setting the stage for an ultimate difference in understanding. The difference in scanpaths can be dangerously different: When novice drivers time share with another in-car task, 40 per cent of the males looked away from the road for over three seconds, an amount of time never seen in the data of experienced drivers (Wikman *et al.* 1998).

The information that gets carried from one processing cycle to another in a dynamic environment is likely to be complexly determined and a function of both the operator's intention and the environment's ability to capture attention. Consider the study by Moray and Rotenberg (1989). Their participants were monitoring a simulated thermal hydraulic system. Of interest here is what happened when a second fault occurred shortly after a first. Normally, operators begin to address a fault within seconds, but because they were processing the first fault, the second fault was not addressed for a considerable time. In fact, the first fault resulted in a drop in attention to all the other subsystems except the one with the initial fault. Moray (1990) notes that the operators occasionally looked at the subsystem containing the second fault, but failed to do anything about it. Integrating the second fault into the eventbase would be difficult if all information beyond the first fault subsystem were pruned. Operator intention (to fix the first fault) weights first fault information heavily, increasing the likelihood that it will be carried over, and thus continue to have a heavy influence on spreading activation. Other nodes would drop out, or, at a minimum, would not be selected for carryover.

Consider results from command, control, and communication settings (C3; Wellens & Ergener 1988; Wellens 1993) when multiple threats manifest. Wellens showed that SA was harmed as communication broke down due to time stress. Under high time stress (events distributed over 10 minutes) compared with low stress (events distributed over 40 minutes), the dispatchers were more reactive, attended more to their own monitor, showed deficits in performance and resource allocation, and showed poorer recall. According to the model discussed in the current chapter, the operator would carry over only part of the information – either detail about one emergency at the expense of another, or limited information about multiple emergencies. If the former occurs, the operator understands one of the emergencies but cognitively dismisses the others. If the latter occurs, the operator may carry over some nodes from each emergency, but then not have the information needed to integrate the emergencies and thus would not deeply understand any of the emergencies. Empirically, the dispatcher often allocated nominal resources to all emergencies; thus some node from each emergency makes it to the next cycle, but with little coherence. Thus, understanding and communication suffer.

The Knowledge Level: Going beyond Explicit Information

The Situation Model in Text Comprehension

The textbase contains the semantic information that is denoted by the linguistic content of the text. However, text comprehension involves much more than just a representation of the meaning of the explicit text content itself. In almost all texts, much of the information necessary for comprehension is only implied or is altogether absent from the text. Thus, to understand fully the situations or ideas described in texts, readers must bring a great deal of prior knowledge to bear. Integrating relevant prior knowledge with the semantic information contained in the textbase gives rise to a representation referred to as the *situation model* (e.g., Kintsch 1998; Zwaan & Radvansky 1998).

The nature of the situation model that is constructed for any given text will depend in part on the kinds of prior knowledge that are integrated with the textbase. Many different kinds of knowledge can be involved in text comprehension, from very general world knowledge to very domain-specific knowledge. For example, the nature of the situation model represented for an expository text often depends on the extent to which the reader has prior knowledge about the particular topic discussed in the text. Several studies have demonstrated that readers with high knowledge within a domain construct more complete, coherent situation models than do readers with low domain knowledge (e.g., Bransford & Johnson 1973; Spilich *et al.* 1979; McNamara *et al.* 1996). Importantly, high and low knowledge readers usually differ not only in the amount of domain-relevant knowledge they have, but also in the organization of that information. For example, experts within a domain are more likely to have well-structured and elaborated *mental models,* which here refers to a representation of the causal (and other) relationships between entities and events that are typical across instances of a situation. Mental models can then be integrated with explicit text content to guide the construction of a situation model for the particular situation being described in a text. Mental models and situation models can be thought of as standing in a type–token relationship. In principle, experts and novices may construct the same textbase but still arrive at different situation models due to differences in the prior knowledge they bring to bear, including differences in the completeness, correctness, and coherence of their knowledge.

Orthogonal to the influence of the amount and structure of prior knowledge, the nature of a situation model will also depend on qualitative differences in the kinds of knowledge incorporated into the text representation. As mentioned above, mental models involve knowledge about typical causal relationships between entities and events as well as knowledge about typical spatial relationships between entities, typical temporal relationships between events, typical goals, emotions, and motivations of protagonists, and so on. Thus, situation models are often multidimensional, and the nature of a particular situation model will depend on the extent to which information along one or more of these dimensions is represented. This may depend in part on the kind of situation being described. The spatial dimension may be particularly important when reading a descriptive text (e.g., driving directions) but may be less central when reading other texts (e.g., a romance novel). Although some texts describe static environments, many expository texts and most narratives describe dynamic environments in which the relationships between entities and events change along several dimensions. Thus, different texts will require different kinds of knowledge and will afford different kinds of situation models.

The situation model is the level of representation that is commonly thought to support performance on tasks that require "deep" comprehension, including problem-solving and application. Additionally, the situation model is thought to support predictive inferences during reading (for discussion of the extent to which readers make predictions while reading, see Millis *et al.* 1990; Fincher-Kiefer 1993; Klin *et al.* 1999; Cook *et al.* 2001).

Empirical Evidence from Text Comprehension Research

Several lines of evidence support the claim that situation models are multidimensional (e.g., Bloom *et al.* 1990; Millis *et al.* 1990; Zwaan *et al.* 1995; Zwaan & Radvansky 1998). For example, Zwaan *et al.* (1998) evaluated the processing-load hypothesis, according to which "the fewer indexes that are shared between the current event being processed and other events in the situation model, the more difficult it should be to incorporate that event into the situation model" (p. 201). Zwaan *et al.* coded each clause in their texts with respect to whether that information could be related to the current situation model along several dimensions, including time (if an event occurred in the same time period as the current situation), space (if an event occurred in the same time period as the current situation), causation (if the situation model contained a causal antecedent for the event), reference (if the information referred to an entity in the current situation), and motivation (if an action is consistent with a protagonist's goal). For all dimensions except for space, reading times increased with increases in the number of dimensions along which a connection between the current clause, and the current situation model could not be formed (a follow-up study showed that coherence on the spatial dimension also predicted reading times when readers memorized a map of the location being described in the text before reading).

Other latency measures have also been used to investigate the situational dimensions that individuals monitor while reading (e.g., Glenberg *et al.* 1987; Zwaan 1996; Scott Rich & Taylor 2000). For example, Glenberg *et al.* (1987) developed texts that introduced a character and an object, and the object was then either described as spatially associated with or dissociated from the character (e.g., John put on/took off his sweatshirt and went jogging). After filler sentences (a sentence mentioning the character but not the object), a probe word naming the object (e.g., sweatshirt) was presented for speeded recognition. Reaction times were longer in the dissociated condition than in the associated condition.

The Situation Model in Situation Awareness

Applied researchers have assumed the existence of situation models with properties like time and causation and recognize their applied value. There is evidence of expert knowledge helping organize situations. Stokes *et al.* (1997) had pilots listen to ATC radio communications. Expert pilots recalled twice the number of concept words, but recalled fewer "filler" words than did apprentices. Experts were also asked to "build a mental picture" of the situation and then select from a set of diagrams that best represented the situation. Experts outperformed apprentices in matching the correct diagram with the dialogue. To Stokes *et al.* "[experts] are better able to make practical use of situational schemata to impose form on sensory data in real time" (p. 191).

It is also thought that the temporal and causal properties of a situation model allow users to anticipate the future and to direct subsequent encodings and pattern recognition. In support, Paull and Glencross (1997) conducted a study on baseball players in which they compared batters' ability to anticipate the direction of a pitch. Experts, who presumably had a good situation model, were quicker and more accurate in making predictions about the pitch. Paull and Glenncross point out that the superior knowledge of experts allowed them to have better anticipation and to identify in the visual display cues that were really useful for the task. In Doane and Sohn (2004) novices were especially poor at predicting the result of multiple, meaningfully related control activities, presumably because the novices did not have the internal model needed to generate predictions from the related control activities.

Sometimes having a model of the situation allows anticipation of the future to be immediate and not a matter of choosing among alternatives. In fact, naturalistic decision-making (see Sieck & Klein, Chapter 8, this volume; Zsambok & Klein 1997) has a central tenet

that perception can lead directly to appropriate action. Perceiving the present tells the expert operator about the future.

Having a model of the situation can also help overcome cognitive limits. For example, working memory (WM) limits are well documented, but they do not always manifest. Durso and Gronlund (1999) argued that retrieval structures (i.e., long-term working memory; LTWM), selection of only important perceptual information, "gistification" of verbatim information, and chunking are all knowledge-dependent avenues that can allow the skilled operator to bypass cognitive limitations. Consider Sohn and Doane's (2004) work with apprentice and expert pilots. They obtained measures of domain-independent spatial WM and knowledge-dependent LTWM. WM was measured using a rotation span task, and knowledge was based on delayed recall of meaningfully vs. non-meaningfully related pairs of cockpit displays. SA was measured by asking participants if a goal would be reached in the next five seconds given a presented cockpit configuration. Of interest here was the fact that, in some analyses, experts and apprentices seemed to rely differentially on general domain-specific WM: as reliance on domain-specific knowledge increased, reliance on general WM decreased.

Designing for Situation Awareness

We end our consideration of SA by considering recent research on design. Advances in understanding the process of comprehension can greatly contribute to the design of more efficient artifacts. Distribution of attention across a display depends on both the display and the operator. Changes in the environment by chance or design certainly affect such attention allocation. Indeed, cognitive ergonomists have a good understanding of how to employ factors like color (Remington *et al.* 2000) and position (Barfield *et al.* 1995; Wickens & Carswell 1995) in the design of effective alarms and displays (Williams 2002). The reader is directed to the excellent introduction to display–design principles presented in Wickens *et al.* (2004).

Because the importance of maintaining SA and meta-SA has been widely recognized, numerous studies have looked at the design of tools that aid in building and maintaining them. In particular, Endsley *et al.* (2003) argue that addressing SA in the design phase is the key to achieve user-centered design.

In order to establish some guidelines on how to design for SA, Endsley *et al.* (2003) first identified eight possible factors (e.g., errant mental models) that might prevent one from having good SA. Based on these eight factors, Endsley *et al.* (2003) formulated a series of design principles that should be followed when designing for SA. In particular, she suggested: organizing the information around goals; supporting the different phases that result in SA (perception, comprehension, projection); making the important cues salient; helping reduce operator's uncertainty; being as simple as possible; keeping the operator in the loop; and supporting the building of shared SA when teams are involved. It is important to keep in mind that these guidelines are not effective if not based on the SA requirements of the specific domain.

Results of different studies support the fact that designing systems with the specific aim of facilitating and enhancing SA is effective. For example, Tlauka *et al.* (2000) researched the effects of a dual map aircraft display – presenting both ego-centered (ERF) and world-centered reference (WRF) frames – on situation awareness. The purpose of a dual display of this type is to support the operator in maintaining both the current navigational path, by means of the ERF, and support a more global SA, through the WRF. Results of the study showed that, after a moderate amount of training, both ERF tasks and WRF tasks improved when relying on dual displays. As another example, Van Breda and Veltman (1998) compared the use of perspective displays with the use of conventional plan-like displays in a target acquisition task. The use of perspective radar displays allowed pilots a faster target acquisition, apparently an SA-dependent behavior.

Studies investigating the effect of Highway-In-The-Sky (HITS) suggest that flight path awareness is better maintained when using the HITS than when using conventional instruments (Haskell & Wickens 1993; Wickens & Prevett 1995). Farley *et al.* (2000) designed an air–ground data-link system with the specific aim of enabling pilots and air traffic controllers to share information expected to enhance SA and the resulting decision-making. The results showed that SA of traffic and weather, as measured by performance-based testable responses, improved. Also, more information shared led to a more collaborative interaction among operators and improved safety.

CONCLUSIONS

In this chapter, we have attempted to draw an analogy between understanding dynamic environments and the comprehension literature that has evolved to explain how readers understand text. From this analogy, we compared methodologies and suggested that methodologies useful in illuminating reading comprehension could be adapted to reveal details about the processes required to understand dynamic situations. We also sketched a model of situation comprehension that can be applied to research conducted under the rubric of situation awareness. Central to the situation comprehension model were processes allowing encoding objects and scenes, an eventbase that allowed integration of events, and a situation model that allowed the operator to employ knowledge of the situation, including causal knowledge, to anticipate the future. Finally, it seems to us that modern research on comprehension of dynamic environments is ready to benefit from more detailed models of situation comprehension. Experiments designed to test the proposed analogy between text comprehension and situation comprehension would contribute to the development of such detailed models.

AUTHOR NOTE

The authors are grateful for the thorough reviews from Steve Lewandowsky, who managed the review process, and from his two anonymous reviewers. Thanks also to the graduate students in Texas Tech's *Cognitive Ergonomics* class for their comments and criticisms.

REFERENCES

Abernethy, B., Gill, D. P., Parks, S. L. & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception*, *30*, 233–252.

Abernethy, B., Maxwell, J. P., Jackson, R. C. & Masters, R. S. W. (2007). Cognition and skill in sport. In F. T. Durso, R. Nickerson, S. Dumais, S. Lewandowsky & T. Perfect (eds.), *The Handbook of Applied Cognition* (2nd edn). Chicester: John Wiley & Sons.

Abernethy, B. & Russell, D. G. (1987). Expert–novice differences in an applied selective attention task. *Journal of Sport Psychology*, *9*, 326–345.

Adams, M. J., Tenney, Y. J. & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors*, *37*, 85–104.

Andre, A. D., Wickens, C. D., Boorman, L. & Boschelli, M. M. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. *International Journal of Aviation Psychology*, *1*, 205–218.

Avila, F. & Moreno, F. J. (2003). Visual search strategies elaborated by tennis coaches during execution error detection processes. *Journal of Human Movement Studies*, *44*, 209–224.

Bar, M. & Ullman, S. (1996). Spatial context in recognition. *Perception*, *25*, 343–352.

Barfield, W., Rosenberg, C. & Furness, T. A. (1995). Situation awareness as a function of frame of reference, computer-graphics eyepoint elevation, and geometric field of view. *International Journal of Aviation Psychology*, *5*, 233–256.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (eds.), *Perceptual Organization* (pp. 213–253). Hillsdale, NJ: Lawrence Erlbaum Associates.

Biederman, I. (1988). Aspects and extensions of a theory of human image understanding. In Z. W. Pylyshyn (ed.), *Computational Processes in Human Vision: An Interdisciplinary Perspective* (pp. 370–428). Norwood, NJ: Ablex.

Biederman, I., Mezzanotte, R. J. & Rabinowitz, J. C. (1982). Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.

Bloom, C. P., Fletcher, C. R., van der Broek, P. *et al.* (1990). An online assessment of causal reasoning during comprehension. *Memory and Cognition*, *18*, 65–71.

Bransford, J. D. & Johnson, M. K. (1973). Consideration of some problems of comprehension. In W. G. Chase (ed.), *Visual Information Processing*. New York: Academic Press.

Busquets, A. M., Parrish, R. V., Williams, S. P. & Nold, D. E. (1994). Comparison of pilots' acceptance and spatial awareness when using EFIS vs. pictorial display formats for complex, curved landing approaches. In R. D. Gilson, D. J. Garland & J. M. Koonce (eds.), *Situational Awareness in Complex Systems: Proceedings of a CAHFA Conference* (pp. 139–167). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Carretta, T. R., Perry, D. C. & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *International Journal of Aviation Psychology*, *6*, 21–41.

Casey, S. (1993). *Set Phasers on Stun and other True Tales of Design, Technology, and Human Error.* Santa Barbara, CA: Aegean Publishing.

Chaparro, A., Groff, L., Tabor, K. *et al.* (1999). Maintaining situational awareness: the role of visual attention. In *Proceedings of the Human Factors and Ergonomics Society* (pp. 1343–1347), Santa Monica, CA: Human Factors and Ergonomics Society.

Chapman, P. & Underwood, G. (1998). Visual search of driving situations: danger and experience. *Perception*, *27*, 951–964.

Chiles, J. R. (2002). *Inviting Disaster: Lessons from the Edge of Technology.* New York: HarperCollins.

Cook, A. E., Limber, J. E. & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language*, *44*, 220–234.

Croft, D. G., Banbury, S. P., Butler, L. T. & Berry, D. C. (2004). The role of awareness in situation awareness. In S. Banbury & S. Tremblay (eds.), *A Cognitive Approach to Situation Awareness: Theory and Application* (pp. 82–103). Aldershot: Ashgate.

Crundall, D. E. & Underwood, G. (1998). Effects of experience and processing demands on visual information acquisition in drivers. *Ergonomics*, *41*, 448–458.

De Graef, P., Christiansen, D. & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*, 317–329.

Deighton, C. D. B. (1997). Towards the development of an integrated human factors and engineering evaluation methodology for rotorcraft D/NAW systems. *QinetiQ Report No. DERA/AS/FMC/CR97629/1.0*, Dec. Farnborough: QinetiQ Ltd.

Dennehy, K. (1997). *Cranfield – Situation Awareness Scale, User Manual.* Applied Psychology Unit, College of Aeronautics, Cranfield University, COA report No. 9702, Bedford.
Doane, S. M. & Sohn, Y. W. (2004). Pilot ability to anticipate the consequences of flight actions as a function of expertise. *Human Factors, 46*, 92–103.
Druckman, D. & Bjork, R. A. (1991). *In the Mind's Eye: Enhancing Human Performance.* Washington, DC: National Academy Press.
Durso, F. T., Bleckley, M. K. & Dattel, A. R. (in press). Does SA add to the validity of cognitive tests? *Human Factors.*
Durso, F. T. & Dattel, A. R. (2004). SPAM: the real-time assessment of SA. In S. Banbury & S. Tremblay (eds.), *A Cognitive Approach to Situation Awareness: Theory and Application* (pp. 137–154). Aldershot: Ashgate.
Durso, F. T. & Gronlund, S. D. (1999). Situation awareness. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt *et al.* (eds.), *Handbook of Applied Cognition* (pp. 283–314). New York: John Wiley & Sons.
Durso, F. T., Hackworth, C., Truitt, T. R. *et al.* (1998a). Situation awareness as a predictor of performance in en route air traffic controllers. *Air Traffic Control Quarterly, 6*, 1–20.
Durso, F. T., Truitt, T. R., Hackworth, C. A. *et al.* (1998b). En route operational errors and situation awareness. *International Journal of Aviation Psychology, 8*, 177–194.
Durso, F. T., Truitt, T. R., Hackworth, C. A. *et al.* (1995). Expertise and chess: comparing situation awareness methodologies. In *International Conference on Situation Awareness*, Daytona, FL.
Endsley, M. R. (1990). Predictive utility of an objective measure of situation awareness. *Proceedings of the Human Factors Society, 34*, 41–45.
Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.
Endsley, M. R. & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology, 4*, 241–264.
Endsley, M. R., Bolté, B. & Jones, D. G. (2003). *Designing for Situation Awareness: An Approach to User-Centered Design.* New York: Taylor and Francis.
Farley, T. C., Hansman, R. J. & Amonlirdviman, K. (2000). Shared information between pilots and controllers in tactical air traffic control. *Journal of Guidance, Control and Dynamics, 23*(5), 826–836.
Farrow, D. & Abernethy, B. (2003). Do expertise and the degree of perception-action coupling affect natural anticipatory performance? *Perception, 32*, 1127–1139.
Fincher-Kiefer, R. (1993). The role of predictive inferences in situation model construction. *Discourse Processes, 16*, 99–124.
Fletcher, C. R. (1981). Short-term memory processes in text comprehension. *Journal of Verbal Learning and Verbal Behavior, 20*, 564–574.
Fletcher, C. R., Hummel, J. E. & Marsolek, C. J. (1990). Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 233–240.
Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108*, 316–355.
Garsoffky, B., Schwan, S. & Hesse, F. W. (2002). Viewpoint dependency in the recognition of dynamic scenes. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 1035–1050.
Glanzer, M. & Nolan, S. D. (1986). Memory mechanisms in text comprehension. In *The Psychology of Learning and Motivation* (Vol. 20, pp. 275–317). San Diego, CA: Academic Press.
Glenberg, A. M., Meyer, M. & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69–83.
Goldman, S. R. & Varma, S. (1995). CAPping the construction-integration model of discourse comprehension. In C. A. Weaver III, S. Mannes & C. R. Fletcher (eds.), *Discourse Comprehension: Essays in Honor of Walter Kintsch* (pp. 337–358). Hillsdale, NJ: Lawrence Erlbaum Associates.
Graesser, A. C. & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading, 2*, 247–269.

Graesser, A. C., Hoffman, N. L. & Clark, L. F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior, 19*, 135–151.
Gugerty, L. J. (1997). Situation awareness during driving: explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3*, 42–66.
Gugerty, L. J. & Falzetta, M. (2005). Using an event-detection measure to assess drivers' attention and situation awareness. In *Proceeding of the 49th Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: Human Factors and Ergonomics Society.
Haber, R. N. & Haber, L. (2003). Perception and attention during low-altitude high-speed flight. In P. S. Tsang & M. A. Vidulich (eds.), *Principles and Practice of Aviation Psychology* (pp. 21–68). Mahwah, NJ: Lawrence Erlbaum Associates.
Haberlandt, K., Berian, C. & Sandson, J. (1980). The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior, 19*, 635–650.
Haberlandt, K. F., Graesser, A. C., Schneider, N. J. & Kiely, J. (1986). Effects of task and new arguments on word reading times. *Journal of Memory and Language, 25*, 314–322.
Haskell, I. D. & Wickens, C. D. (1993). Two- and three-dimensional displays for aviation: a theoretical and empirical comparison. *International Journal of Aviation Psychology, 3*, 87–109.
Hauβ, Y., Gauss, B. & Eyferth, K. (2000). The evaluation of a future air traffic management: towards a new approach to measure situation awareness in air traffic control. In L. M. Camarinha-Matos, H. Afsarmanesh & H. Herbe (eds.), *Advances in Networked Enterprises: Virtual Organization, Balanced Automation and System Integration* (CD Suppl.). Boston, MA: Kluwer.
Hauβ, Y., Gauss, B. & Eyferth, K. (2001). The influence of multi-sector planning on the controllers' mental models. In D. Harris (ed.), *Engineering Psychology and Cognitive Ergonomics* (pp. 203–209). Aldershot: Ashgate.
Helsen, W. F. & Pauwels, J. M. (1993). The relationship between expertise and visual information processing in sport. *Applied Cognitive Psychology, 13*, 109–134.
Helsen, W. F. & Starkes, J. L. (1999). A multidimensional approach to skilled perception and performance in sport. *Applied Cognitive Psychology, 13*, 1–27.
Horswill, M. S. & McKenna, F. P. (2004). Drivers' hazard perception ability: situation awareness on the road. In S. Banbury & S. Tremblay (eds.), *A Cognitive Approach to Situation Awareness: Theory and Application.* Aldershot: Ashgate.
Howarth, C., Walsh, W. D. & Abernethy, B. (1984). A field examination of anticipation in squash: some preliminary data. *The Australian Journal of Science and Medicine in Sport, 16*, 7–11.
James, N. & Patrick, J. (2004). The role of situation awareness in sport. In S. Banbury & S. Tremblay (eds.), *A Cognitive Approach to Situation Awareness: Theory and Application* (pp. 137–154). Aldershot: Ashgate.
Jeannot, E., Kelly, C. & Thompson, D. (2003). *The Development of Situation Awareness Measures in ATM Systems.* HRS/HSP-005-REP-01. Brussels: EUROCONTROL.
Jensen, S. (1999). Perceived versus real situation awareness: towards more objective assessment of SA. In L. Straker & C. Pollack (eds.), *CD-ROM Proceedings of CybErg 1999: The Second International Cyberspace Conference on Ergonomics* (pp. 327–334). Perth: International Ergonomics Association Press.
Jones, D. G. & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine, 67*, 507–512.
Jones, D. G. & Kaber, D. B. (2004). Situation awareness measurement and the situation awareness global assessment technique. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas & H. Hendrick (eds.), *Handbook of Human Factors and Ergonomics Methods* (pp. 42-1–42-8). Boca Raton, FL: CRC Press.
Just, M. A. & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*, 329–354.
Kendeou, P. & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology, 97*, 235–245.
Kintsch, W. (1988). The use of knowledge in discourse processing: a construction-integration model. *Psychological Review, 95*, 163–182.
Kintsch, W. (1994). Learning from text. *American Psychologist, 49*, 294–303.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.
Kintsch, W., Kozminsky, E., Streby, W. J. *et al.* (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior, 14*, 196–214.
Kintsch, W., Welsch, D. M., Schmalhofer, F. & Zimny, S. (1990). Sentence memory: a theoretical analysis. *Journal of Memory and Language, 29*, 133–159.
Klin, C. M., Murray, J. D., Levine, W. H. & Guzmán, A. E. (1999). Forward inferences: from activation to long-term memory. *Discourse Processes, 27*, 241–260.
Lachman, R., Lachman, J. L. & Butterfield, E. C. (1979). *Cognitive Psychology and Information Processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Long, D. L., Golding, J. M. & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language, 31*, 634–647.
McCarley, J. S., Kramer, A. F., Wickens, C. D. *et al.* (2004). Visual skills in airport security screening. *Psychological Science, 15*, 302–306.
McKenna, F. P. & Crick, J. L. (1991). *Hazard Perception in Drivers: A Methodology for Testing and Training*. Final Report. Behavioral Studies Unit, Transport and Road Research Laboratory, Crowthorne, UK.
McKoon, G., Gerrig, R. J. & Greene, S. B. (1996). Pronoun resolution without pronouns: some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 919–932.
McKoon, G., Ratcliff, R., Ward, G. & Sproat, R. (1993). Syntactic prominence effects on discourse processes. *Journal of Memory and Language, 32*, 593–607.
McNamara, D. S., Kintsch, E., Songer, N. B. & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1–43.
McNamara, D. S. & Kintsch, W. (1996). Learning from texts: effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247–288.
Magliano, J. P., Graesser, A. C., Eymard, L. A. *et al.* (1993). Locus of interpretive and inference processes during text comprehension: a comparison of gaze durations and word reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 704–709.
Magliano, J. P., Trabasso, T. & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*, 615–629.
Maki, R. H. (1998). Test predictions over text material. In D. J. Hacker, J. Dunlosky & A. C. Graesser (eds.), *Metacognition in Educational Theory and Practice* (pp. 117–144). Hillsdale, NJ: Lawrence Erlbaum Associates.
Maki, R. H., Foley, J. M., Kajer, W. K. *et al.* (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory & Cognition, 16*, 609–616.
Maki, R. H., Jonas, D. & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review, 1*, 126–129.
Malt, B. C. (1985). The role of discourse structure in understanding anaphora. *Journal of Memory and Language, 24*, 271–289.
Mayer, R. E. & Jackson, J. (2005). The case for coherence in scientific explanations: qualitative details can hurt qualitative understanding. *Journal of Experimental Psychology: Applied, 11*, 13–18.
Mayer, R. E., Bove, W., Bryman, A. *et al.* (1996). When less is more: meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*, 64–73.
Metalis, S. A. (1993). Assessment of pilot situational awareness: measurement via simulation. In *Proceedings of the Human Factors Society 37th Annual Meeting* (pp. 113–117). Santa Monica, CA: The Human Factors and Ergonomics Society.
Millis, K. K., Morgan, D. & Graesser, A. C. (1990). The influence of knowledge-based inferences on reading time for expository text. In A. C. Graesser & G. H. Bowers (eds.), *The Psychology of Learning and Motivation: Inferences and Text Comprehension*. New York: Academic Press.

Millis, K. K., Simon, S. & tenBroek, N. S. (1998). Resource allocation during the rereading of scientific texts. *Memory & Cognition, 26*, 232–246.
Mills, K. L., Hall, R. D., McDonald, M. & Rolls, G. W. P. (1998). *The Effects of Hazard Perception Training on the Development of Novice Driver Skills*. Research Report. London: Department for Transport.
Moray, N. (1990). Designing for transportation safety in the light of perception, attention, and mental models. *Ergonomics, 33*, 1201–1213.
Moray, N. & Rotenberg, I. (1989). Fault management in process control: eye movements and action. *Ergonomics, 32*, 1319–1342.
Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Memory & Cognition, 12*, 489–497.
Myers, J. L., Shinjo, M. & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*, 453–465.
Neal, A., Griffin, M., Paterson, J. & Bordia, P. (1998). *Human Factors Issues: Performance Management Transition to a CNS/ATM Environment*. Final Report: Air Services Australia. Brisbane: University of Queensland.
O'Brien, E. J. & Albrecht, J. E. (1991). The role of context in accessing antecedents in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 94–102.
O'Brien, E. J., Raney, G. E., Albrecht, J. E. & Rayner, K. (1997). Processes involved in the resolution of explicit anaphors. *Discourse Processes, 23*, 1–24.
O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*, 540–552.
Oliva, A. & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34*, 72–107.
Oliva, A. & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*, 176–210.
Oppenheimer, R. (1956). Analogy in science. *American Psychologist, 11*, 127–135.
Paull, G. & Glencross, D. (1997). Expert perception and decision making in baseball. *International Journal of Sport Psychology, 28*, 35–56.
Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning & Memory, 2*, 509–522.
Rawson, K. A. & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 69–80.
Rawson, K. A., Dunlosky, J. & McDonald, S. L. (2002). Influences of metamemory on performance predictions for text. *Quarterly Journal of Experimental Psychology, 55A*, 505–524.
Rawson, K. A. & Kintsch, W. (2004). Exploring encoding and retrieval effects of background information on text memory. *Discourse Processes, 38*, 323–344.
Rawson, K. A. & Kintsch, W. (2005). Rereading effects depend upon time of test. *Journal of Educational Psychology, 97*, 70–80.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Rayner, K., Kambe, G. & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology, 53A*, 1061–1080.
Remington, R. W., Johnston, J. C., Ruthruff, E. *et al.* (2000). Visual search in complex displays: factors affecting conflict detection by air traffic controllers. *Human Factors, 42*, 349–366.
Renninger, L. W. & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research, 44*, 2301–2311.
Roeder, K. D. (1962). The behaviour of free flying moths in the presence of artificial ultrasonic pulses. *Animal Behaviour, 10*, 300–304.
Roscoe, S. N. (1997). *Predicting Human Performance*. Quebec: Helio Press.
Rousseau, R., Tremblay, S. & Breton, R. (2004). Defining and modeling situation awareness: a critical review. In S. Banbury & S. Tremblay (eds.), *A Cognitive Approach to Situation Awareness: Theory and Application* (pp. 3–21). Aldershot: Ashgate.
Salas, E., Cannon-Bowers, J. A., Fiore, S. M. & Stout, R. J. (2001). Cue-recognition training to enhance team situation awareness. In M. McNeese & E. Salas (eds.), *New Trends in Cooperative*

Activities: Understanding System Dynamics in Complex Environments. Santa Monica, CA: Human Factors and Ergonomics Society.

Sarter, N. B. & Woods, D. D. (1991). Situation awareness: a critical but ill-defined phenomenon. *International Journal of Aviation Psychology, 1*, 45–57.

Scott Rich, S. & Taylor, H. A. (2000). Not all narrative shifts function equally. *Memory & Cognition, 28*, 1257–1266.

Schyns, P. G. & Oliva, A. (1994). From blobs to boundary edges: evidence for time and spatial scale dependent scene recognition. *Psychological Science, 5*, 195–200.

Secrist, G. E. & Hartman, B. O. (1993). Situational awareness: the trainability of the near-threshold information acquisition dimension. *Aviation, Space, and Environmental Medicine, 64*, 885–892.

Sieck, W. R. & Klein, G. (2007). A comparison of rationalist and naturalistic decision methods for applied decision research. In F. T. Durso, R. Nickerson, S. Dumais *et al.* (eds.), *The Handbook of Applied Cognition* (2nd edn). Chichester: John Wiley & Sons.

Singer, M., Halldorson, M., Lear, J. C. & Andrusiak, P. (1992). Validation of causal bridging inference in discourse understanding. *Journal of Memory and Language, 31*, 507–524.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L. & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18*, 275–290.

Sohn, Y. W. & Doane, S. M. (2004). Memory processes of flight situation awareness: interactive Roles of working memory capacity, long-term working memory, and expertise. *Human Factors, 46*, 461–475.

Stokes, A., Kemper, K. & Kite, K. (1997). Aeronautical decision-making, cue recognition, and expertise under time pressure. In C. E. Zsambok & G. Klein (eds.), *Naturalistic Decision Making* (pp. 183–197). Mahwah, NJ: Lawrence Erlbaum Associates.

Suh, S. & Trabasso, T. (1993). Inferences during reading: converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32*, 279–300.

Taylor, R. M. (1990). Situational awareness rating technique (SART): the development of a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations* (AGARD-CP-478; pp. 3/1–3/17). Neuilly-sur-Seine: NATO-AGARD.

Thiede, K. W. & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*, 129–160.

Tlauka, M., Stanton, D. & McKenna, F. P. (2000). Dual displays. *Ergonomics, 43*(6), 764–770.

Underwood, G., Chapman, P., Bowden, K. & Crundall, D. (2002). Visual search while driving: skill and awareness during inspection of the scene. *Transportation Research Part F, 5*, 87–97.

Underwood, G., Chapman, P., Brocklehurst, N. *et al.* (2003). Visual attention while driving: sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*, 629–646.

Underwood, G., Crundall, D. & Chapman, P. (2007). Cognition and driving. In F. T. Durso, R. Nickerson, S. Dumais, S. Lewandowsky & T. Perfect (eds.), *The Handbook of Applied Cognition* (2nd edn). Chicester: John Wiley & Sons.

Van Breda, L. & Veltman, H. A. (1998). Perspective information in the cockpit as target acquisition aid. *Journal of experimental psychology: Applied, 4*, 55–68.

Vidulich, M. A. & Hughes, E. R. (1991). Testing a subjective metric of situation awareness. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1307–1311). Santa Monica, CA: The Human Factors and Ergonomics Society.

Vidulich, M. A., Stratton, M., Crabtree, M. & Wilson, G. (1994). Performance-based and physiological measures of situational awareness. *Aviation, Space, and Environmental Medicine, 65*, A7–A12.

Waag, W. L. & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space, and Environmental Medicine, 65*, A13–A19.

Wellens, A. R. (1993). Group situation awareness and distributed decision making: from military to civilian applications. In N. J. Castellan (ed.), *Individual Group and Decision Making: Current Issues* (pp. 267–291). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wellens, A. R. & Ergener, D. (1988). The C.I.T.I.E.S. game: A computer-based situation assessment task for studying distributed decision-making. *Simulation & Games*, 19, 304–327.

Wickens, C. D. & Carswell, C. M. (1995). The proximity compatibility principle: its psychological foundation and relevance to display design. *Human Factors, 37*, 473–494.

Wickens, C. D., Lee, J., Liu, Y. D. & Becker, S. (eds.) (2004). *An Introduction to Human Factors Engineering.* Upper Saddle River, NJ: Pearson Education.

Wickens, C. D. & Prevett, T. T. (1995). Exploring the dimension of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied, 1*, 110–135.

Wiggins, M. & O'Hare, D. (2003). Weatherwise: evaluation of a cue-based training approach for the recognition of deteriorating weather conditions during flight. *Human Factors, 45*, 337–345.

Wikman, A. S., Nieminen, T. & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics, 41*, 358–372.

Wiley, J., Mason, R. A. & Myers, J. L. (2001). Accessibility of potential referents following categorical anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1238–1249.

Wiley, J. & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: evidence from eye movements. *Memory & Cognition, 28*, 1011–1021.

Williams, A. M. & Davids, K. (1998). Visual search strategy, selective attention and expertise in soccer. *Research Quarterly for Exercise and Sport, 69*, 111–128.

Williams, A. M., Ward, P., Knowles, J. M. & Smeeton, N. J. (2002). Anticipation skill in a real-world task: measurement, training and transfer in tennis. *Journal of Experimental Psychology: Applied, 8*, 259–270.

Williams, K. W. (2002). Impact of highway-in-the-sky displays on pilot situation awareness. *Human Factors, 44*, 18–27.

Woodhouse, R. & Woodhouse, R. A. (1995). Navigation errors in relation to controlled flight into terrain (CFIT) accidents. In *8th International Symposium on Aviation Psychology.* Columbus, OH, April.

Woods, D. (1991). The cognitive engineering of problem representations. In J. Alty & G. Weir (eds.), *Human–Computer Interaction and Complex Systems* (pp. 169–188). London: Academic Press.

Yntema, D. B. (1963). Keeping track of several things at once. Human Factors, *5*, 7–17.

Yntema, D. B. & Mueser, G. E. (1960). Remembering the present states of a number of variables. *Journal of Experimental Psychology, 60*, 18–22.

Yntema, D. B. & Mueser, G. E. (1962). Keeping track of variables that have few or many states. *Journal of Experimental Psychology, 63*, 391–395.

Zsambok, C. E. & Klein, G. (eds.) (1997). *Naturalistic Decision Making. Expertise: Research and Applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 920–933.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1196–1207.

Zwaan, R. A., Magliano, J. P. & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386–397.

Zwaan, R. A. & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E. & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading, 2*, 199–220.